



(19)

BEST AVAILABLE COPY

(11) Publication number:

10078969 A

Generated Document

PATENT ABSTRACTS OF JAPAN

(21) Application number: 08233485

(51) Intl. Cl.: G06F 17/30

(22) Application date: 03.09.96

(30) Priority:

(43) Date of application
publication: 24.03.98(84) Designated
contracting states:(71) Applicant: NIPPON TELEGR & TELEPH CORP
<NTT>(72) Inventor: BESSHO KATSUTO
IWASE SHIGETO

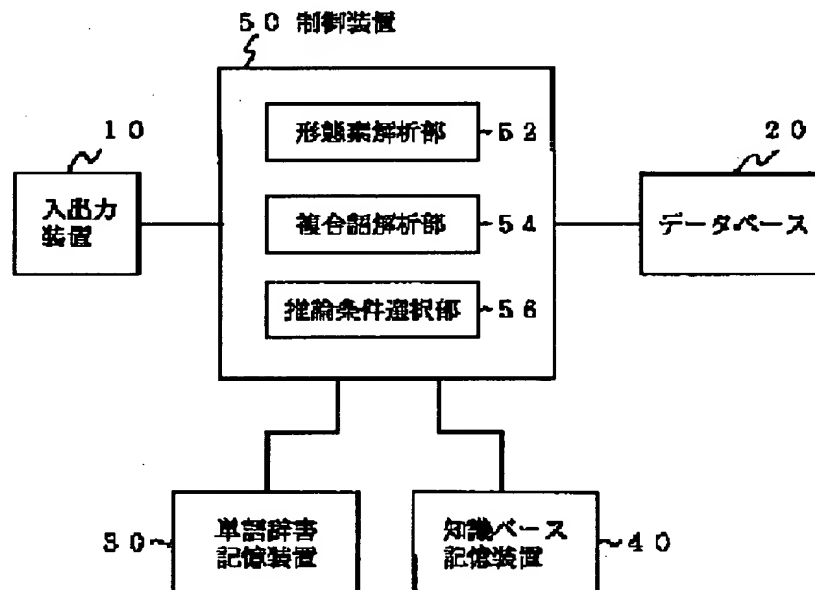
(74) Representative:

(54) INFORMATION
RETRIEVING DEVICE

(57) Abstract:

PROBLEM TO BE SOLVED: To provide an information retrieving device which can retrieve an input inquiry statement of versatile compound words.

SOLUTION: An input/output device 10 inputs an inquiry statement in natural sentences and outputs the information accordant with the input condition. A data base 20 stores the retrieval object information including the names, products, occupations, etc. A word dictionary 30 stores the information on words constructing a keyword included in the base 20. A knowledge base 40 stores the sorting of occupations, the combinations of products and verbs and various thesauruses. A compound word analysis part 54 analyzes a keyword of compound words, deletes the front parts, the ending of words and the 'SA'-irregular conjugation nouns, and separates the products from the verbs. An inference condition selection part 56 selects a combination of a product and a verb necessary for the inference from plural types of keywords. Then a controller 50 infers an occupation name, etc., from the product and the verb that are finally decided.



COPYRIGHT: (C)1998,JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-78969

(43) 公開日 平成10年(1998) 3月24日

(51) Int.Cl.⁸

G 0 6 F 17/30

識別記号

庁内整理番号

F I

G 0 6 F 15/403

3 3 0 C

3 2 0 D

技術表示箇所

審査請求 未請求 請求項の数 6 O L (全 7 頁)

(21) 出願番号 特願平8-233485

(22) 出願日 平成8年(1996) 9月3日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 別所 克人

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 岩瀬 成人

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

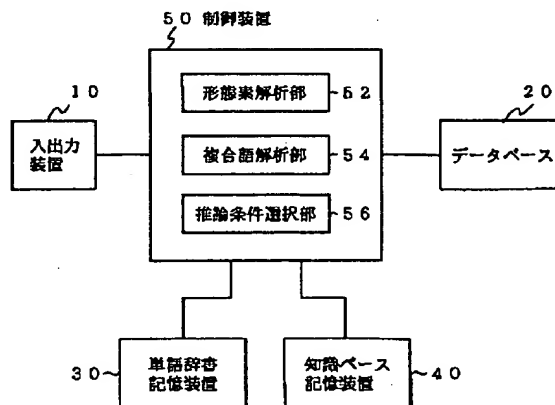
(74) 代理人 弁理士 鈴木 誠

(54) 【発明の名称】 情報検索装置

(57) 【要約】

【課題】 多様な複合語の入力問合せ文に対応する検索を可能にする。

【解決手段】 入出力装置10は自然文で問合せ文を入力し、入力条件に合致する情報を出力する。データベース20は名義、商品、職業等を含む検索対象情報を格納する。単語辞書30はデータベース中のキーワードを構成する単語情報を記憶する。知識ベース40は、データベースの職業分類と商品及び動詞の組合せ、更に、種々のシソーラスを記憶する。複合語解析部54は、複合語のキーワードを解析し、前方削除や語尾・サ変名詞削除をしたり、商品と動詞に分離したりする。推論条件選択部56は、複数種類のキーワードについて、それらのキーワードから推論に必要な商品と動詞の組合せを選択する。制御装置50は、このようにして最終的に決定された商品と動詞を用いて職業名等を推論する。



【特許請求の範囲】

【請求項1】 自然文で問合せ文を入力し、入力条件に合致する情報を出力する情報検索装置において、名義、商品、職業等を含む情報を格納するデータベースと、

前記データベース中の名義、商品、職業等のキーワードを構成する単語と各単語の属性を対の形で記憶しておく単語辞書記憶装置と、

前記データベースの職業分類と、該職業分類の業務対象を商品と動詞の組合せて記憶し、さらに商品や動詞シソーラスを記憶する知識ベース記憶装置と、

複合語のキーワードが入力されたとき該複合語を解析し、商品と動詞に分離する複合語解析手段と、複数の種類のキーワードから推論に使用する商品と動詞を選択する推論条件選択手段を具備し、

前記複合語解析手段は、複合語のキーワードを前方削除することにより、前記知識ベース記憶装置中のシソーラスを再検索することを特徴とする情報検索装置。

【請求項2】 前記複合語解析手段は、商品キーワードが語尾やサ変名詞を含む場合、該語尾やサ変名詞を削除して前記知識ベース記憶装置中の商品シソーラスを再検索し、削除した語尾やサ変名詞を動詞化して前記知識ベース記憶装置中の動詞シソーラスを検索することを特徴とする請求項1記載の情報検索装置。

【請求項3】 前記複合語解析手段は、職業キーワードが複合語の場合、該職業キーワードを商品部とサ変名詞・職業語尾等の動詞部に分離し、商品部で商品シソーラスを検索し、動詞部は動詞化することにより動詞シソーラスを検索することを特徴とする請求項1記載の情報検索装置。

【請求項4】 前記知識ベース記憶装置は商品を別の商品と動詞の組合せて表わした事象変換テーブルを具備し、前記複合語解析手段は、該事象変換テーブルを用いて、商品キーワードまたは職業キーワードから抽出した商品と動詞の組合せに対して、商品を動詞に対応した商品に変換することを特徴とする請求項1記載の情報検索装置。

【請求項5】 前記知識ベース記憶装置は、商品を別の商品と動詞の組合せて表わした事象変換テーブル、及び、動詞のとるべき商品の種別と動詞の組合せと、当該組合せから変換可能な商品種別と動詞の組合せを表わすプランニング知識テーブルを具備し、

前記複合語解析手段は、商品キーワード、動詞キーワード、職業キーワードから得られた商品と動詞に対し、その商品の商品種別と動詞から動詞を前記プランニング知識テーブルを参照して別の動詞に変換し、さらに前記事象変換テーブルを参照して、商品を新しく得られた動詞に対応した商品に変換することを特徴とする請求項1記載の情報検索装置。

【請求項6】 前記推論条件手段は、商品キーワードと

職業キーワード及び動詞キーワードが入力されたときは、入力された商品キーワードと商品キーワード中の商品部、商品キーワードと職業キーワード中の動詞部と動詞キーワードから、商品と動詞を組合せてそれぞれ推論することを特徴とする請求項1記載の情報検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報検索装置に係り、詳しくは、イエローページサービス等の職業情報案内サービスにおいて、問合せがデータベースに登録されている職業名と異なっていたり、問合せ中に職業名が含まれていなくても検索を可能とする情報検索装置に関するものである。

【0002】

【従来の技術】従来の自然語による情報検索では、検索に使用するためのキーワードが一語からなる単語または2語程度からなる複合語を対象としており、複合語のままシソーラスに登録し、検索処理を行っている。また、文献検索の分野では、複合語のキーワードを単語に分解し、それぞれの単語で検索を行っている。

【0003】一方、電話帳等の職業情報案内サービスにおいて、例えば特開昭64-29929号公報に記載のように、問合せ文中に職業名がなくても問合せ文中の商品情報と動詞情報により職業を推論して、データベースを検索可能とするものもある。

【0004】

【発明が解決しようとする課題】あらゆる複合語のキーワードをシソーラスに登録することは、複合語を構成する単語の組合せの数が膨大になるので、非現実的である。そこで、シソーラスに登録される単語は、1～2語程度の単語に限定される。従来の自然語による情報検索では、入力されたキーワードをそのまま使用してシソーラスを検索するので、検索に使用できるキーワードに限定があり、ユーザの真の意図に合致する検索結果が得られないという問題があった。

【0005】また、複合語を単語に分解し、それぞれの単語で検索する方法では、問合せと無関係な結果が得られるという問題がある。特に職業情報案内の分野では適切な職業のみ推論することが重要である。例えば複合語中の単語の役割を意識しないと、「競馬情報」という入力力で「情報」から「情報科学」を検索してしまうという問題がある。さらに、「布団を扱うレンタルショップ」の様に動詞キーワードが職業（「レンタルショップ」）の中に含まれている場合がある。この場合は動詞キーワード「扱う」から推論に必要な動詞「レンタルする」を導き出すことが出来ない。

【0006】更に、特開昭64-29929号公報に記載の方法は、単純な意味解析によって問合せ文中から商品情報と動詞情報を得るものであるため、多様な複合語の入力キーワードに対応できない問題があった。

【0007】本発明の目的は、シソーラスには1～2語程度の要素からなる単語を登録しておくだけで、多様な複合語の入力キーワードに対応することが可能な情報検索装置を提供することにある。

【0008】本発明の他の目的は、複数の種類のキーワードが入力された場合でも、それらのキーワードから推論に必要な商品と動詞の組合せを選択して推論することで、正確に検索することが可能な情報検索装置を提供することにある。

【0009】

【課題を解決するための手段】複合語の商品キーワードされた時、語尾（複合語の末尾にあって、複合語の中では主要な意味をもたない単語）やサ変名詞（「する」をつけて動詞になる名詞。例：「予想」）が含まれるときは、語尾やサ変名詞を動詞化し、商品キーワードからは語尾やサ変名詞は削除する。商品キーワードに語尾やサ変名詞が含まれないときは、商品キーワードを前方削除してキーワードを再構成する。

【0010】複合語の職業キーワードが入力されたときは、職業キーワードを商品部分と、語尾サ変等の動詞部分に分離し、商品部分から商品キーワード、動詞部分から動詞キーワードを再構成する。

【0011】このようにして、商品キーワードと動詞キーワードを作り直し、知識ベース記憶装置中のシソーラスを検索する。商品キーワードまたは動詞キーワードからシソーラス上での商品と動詞が決定すれば、商品と動詞に対応した種類に変換する。

【0012】以上の過程から決定した商品と動詞を用いて、データベースの職業分類と、該職業分類の業務対象を商品と動詞の組合せで記憶する知識ベースを検索し、商品と動詞から職業を推論する。また、商品キーワードと職業キーワード及び動詞キーワードが入力されたときは、複合語解析装置が出力した、商品キーワード・職業キーワード中の商品、動詞と動詞キーワードを組合せて推論する。さらに、商品と動詞の組合せに対する職業分類がない時、その商品と動詞の組合せを、その意味を損ねないような商品と動詞の組合せに変換して関連を再検索する。

【0013】一般に日本語では複合語を構成する単語の内、後方の部分に重要な意味をもっているものが多いため、前方削除によりシソーラス中の単語を検索することにより適切な商品を検索できる。一方、日本語では、複合語の末尾にくる単語があまり意味のない単語であったり、サ変名詞のように動詞化できることもあり、そのような末尾語やサ変名詞を動詞化し、末尾語・サ変名詞を除いた複合語中の意味的に重要な部分でシソーラス中の単語を検索することにより適切な商品を検索できる。

【0014】商品キーワードや職業キーワードから抽出した商品と動詞は、商品が動詞のとり格に設定できない場合があるので、商品と動詞の格に対応した商品に変換

することにより適切な商品と動詞の組を得ることができる。また、複数の種類のキーワードがあった場合でも、それらのキーワードから推論に必要な商品と動詞の組合せを選択して推論するので、検索結果に誤りをなくすることができる。

【0015】

【発明の実施の形態】以下、本発明の一実施例について図面により説明する。

【0016】図1は本発明における情報検索装置の一実施例の全体構成図を示す。図において、10は問合せ文を入力し、検索結果を出力する入出力装置、20は名義、商品、職業等を含む情報を格納するデータベース、30は名義、商品、職業等のキーワードを構成する単語と各単語の属性を対の形で記憶しておく単語辞書記憶装置、40は種々の知識を記憶する知識ベース記憶装置、50は全体を制御する制御装置である。ここで、制御装置50は、本発明に関連する構成として、入力された問合せ文を単語辞書記憶装置30を用いて単語単位に分解する形態素解析部52、複合語の商品キーワードや職業キーワードなどを解析し、商品と動詞に分離する複合語解析部54、複数の種類のキーワードから推論に使用する商品と動詞を選択する推論条件選択部56を有する。

【0017】図2は、知識ベース記憶装置40に記憶される知識の一覧である。ここで、データベースの職業と、該職業分類の業務対象を商品と動詞の組合せで表わした集合（職業分類＝商品＋動詞）は、最終的に商品と動詞から職業を推論するときに使用される。語尾動詞変換テーブルは職業語尾と動詞の対応を表わしたテーブルである（例：音楽教室の教室→教える）。事象変換テーブルは商品と動詞の組合せで表わしたテーブルである（例：調理師→調理する）。ブランチング知識テーブルは、動詞のとりべき商品の種別と動詞の組合せと、その組合せから変換可能な商品種別と動詞の組合せを表わしたテーブルである（例：（資格）取る→（事）習う）。

【0018】以下では、入出力装置10から入力された問合せ文は、すでに形態素解析部52において単語辞書記憶装置30を用いて単語単位に分解されているとして、特に複合語解析部54の動作について詳述する。

【0019】まず、図3を用いて、複合語解析部54が複合語の商品キーワードからシソーラス中の商品と動詞を求める処理例を説明する。

【0020】ステップ110；まず、商品キーワード全体で、知識ベース記憶装置40の商品シソーラスを検索する。ここで、検索に成功すれば、それを商品として終了する。失敗した場合、ステップ120に進む。

【0021】ステップ120；末尾語が語尾またはサ変名詞か判断する。語尾またはサ変名詞ならば、ステップ130に進む。それ以外ならば、ステップ140に進む。例えば、「競馬予想」「競馬情報」「ピアノの先

生」では、それぞれ「予想」はサ変名詞、「情報」「先生」は語尾なので、ステップ130に進む。また、「家庭菜園」の末尾語「菜園」は語尾でもサ変名詞でないで、ステップ140に進む。

【0022】ステップ130；末尾語が語尾の場合は、知識ベース記憶装置40の語尾動詞変換テーブルを参照し、動詞に変換する。サ変名詞の場合は、末尾に「する」を付加して動詞にする。例えば、「予想」は「予想する」、「先生」は「教える」にする。「情報」は動詞には対応しないので、動詞なしとする。その後、末尾語を除いた部分で知識ベース記憶装置40の商品シソーラスを検索する。例えば、「競馬予想」の場合は「競馬」で商品シソーラスを検索する。そして、成功すれば、それを商品とし終了する。失敗した場合は、末尾語を除いたままステップ120に戻り、次に、末尾語を除いた部分の末尾語を判定する。

【0023】ステップ140；それまで削除した末尾語を全て戻した後、先頭語がサ変名詞相当かどうか判断する。そして、サ変名詞相当ならば、ステップ150に進む。そうでなければ、ステップ160に進む。例えば、「レンタルCD」「貸し衣装」の「レンタル」や「貸し」はサ変名詞相当で、それぞれ「レンタルする」や「貸す」に動詞化できるので、ステップ150に進む。

【0024】ステップ150；先頭語を動詞化する。その後、先頭語を除いた部分で知識ベース記憶装置40の商品シソーラスを検索する。例えば、「レンタルCD」の場合は「CD」で検索する。そして、成功すれば、それを商品とし終了する。失敗した場合は、先頭語を除いたままステップ120に戻る。

【0025】ステップ160；先頭語を削除して、知識ベース記憶装置40の商品シソーラスを検索する。例えば、「ビジネス英語」の場合「ビジネス」を削除し、「英語」で検索する。そして、検索に成功すれば、それを商品として、終了する。失敗した場合は、先頭語を削除したままステップ120に戻る。

【0026】図4に、複合語の商品キーワードから商品と動詞を求める処理における商品キーワードの遷移過程を示す。即ち、まず、末尾語にくる語尾やサ変名詞を削除していき、末尾語の語尾やサ変名詞がなくなったならば、一旦削除した語尾・サ変名詞を全てつけ戻した上で、次に、先頭語を削除し、再び末尾の語尾・サ変名詞を削除していく。図4は、最終的に、ステップ130で単語「C」が検索に成功したことを示している。

【0027】次に、図5を用いて、複合語解析部54が複合語の職業キーワードからシソーラス中の商品と、動詞を求める処理例を説明する。

【0028】ステップ210；複合語の職業キーワードを一語ずつ前方削除しながら、知識ベース記憶装置40の職業シソーラスを検索する。例えば、「温泉旅館」の場合、「温泉」を削除して、「旅館」で検索する。そし

て、検索に成功すれば、それを商品として、終了する。失敗した場合は、ステップ220に進む。

【0029】ステップ220；職業語尾（職業相当の語尾）があるかどうか判断する。そして、職業語尾があるならば、ステップ230に進む。なければ終了する。例えば、「音楽教室」では「教室」は職業語尾なので、ステップ230に進む。「レストラン」は職業語尾でないで終了する。

【0030】ステップ230；まず、職業語尾以外の語を商品キーワードとし、職業語尾は知識ベース記憶装置40の語尾変換テーブルを参照して動詞に変換し動詞キーワードとする。例えば、「音楽教室」では「教室」を商品キーワードに、「教室」を「教える」に動詞化して動詞キーワードにする。次に、商品キーワードについて、図3で述べた処理を行い、シソーラス中の商品と動詞を求める。

【0031】次に、図6を用いて、複合語解析部54が職業キーワードから動詞を抽出する処理例を説明する。

【0032】ステップ310；職業語尾の直前の単語がサ変名詞相当の単語かどうか判断する。職業語尾がない場合は、末尾の語についてサ変名詞相当かどうか判断する。そして、サ変名詞相当の単語ならば、ステップ320に進む。そうでなければ、ステップ330に進む。例えば、「ビデオレンタル店」では、「レンタル」はサ変名詞なので、ステップ320に進む。

【0033】ステップ320；サ変名詞相当の単語を動詞化して終了する。

【0034】ステップ330；職業語尾があるかどうか判断する。あれば、ステップ340に進む。なければ、ステップ350に進む。例えば、「音楽教室」では「教室」は職業語尾なので、ステップ340に進む。

【0035】ステップ340；職業語尾を知識ベース記憶装置40の語尾動詞変換テーブルを参照して動詞に変換して終了する。

【0036】ステップ350；職業キーワードの先頭の単語からサ変名詞相当の単語かどうか調べていく。サ変名詞相当の単語が見つければ、ステップ360に進む。見つからなければ、終了する。例えば、「貸しビデオ店」では、「貸し」はサ変名詞相当の単語なので、ステップ360に進む。

【0037】ステップ360；サ変名詞相当の単語を動詞化して終了する。例えば、サ変名詞相当の単語「貸し」は「貸す」に動詞化して終了する。

【0038】次に、図7を用いて、複合語解析部54が商品キーワードまたは職業キーワードから抽出した商品と動詞の組合せに対して、商品を動詞に対応した商品に変換する商品最適変換の処理例を説明する。

【0039】知識ベース記憶装置40においては、各商品には物、事、人等の商品種別が割り当てられている。例えば、「布団」、「競馬」「調理師」の商品種別はそ

れぞれ「物」、「事」、「人」である。一方、動詞にはその動詞がとるべき商品の種別が定まっている。例えば、「教える」がとるべき商品の種別は「事」である。

【0040】キーワードから商品と動詞を抽出した場合、動詞と商品種別との不整合がある場合がある。このような場合、複合解析部54では、図7のように、知識ベース記憶装置40の商品を動詞に対応した商品種類の商品に対応づけしている事象変換テーブル46を用いて商品の最適変換を行う。

【0041】例えば、職業キーワードが「調理師／専門学校」の場合、職業中商品として「調理師（商品種類：人）」が抽出され、職業中動詞として「教える」（専門学校→教える）が語尾動詞変換テーブル45により抽出される。しかしながら、「教える」がとるべき商品種別は「事」であるので、「調理師（人）」が「する事」は何かを、事象変換テーブル46を参照して変換すると「調理（事）」という知識が得られる。そこで、商品を「調理」に変換する。これによって、職業分類および動詞との組合せ集合により、商品「調理」と動詞「教える」から職業「料理学校」を推論することができる。

【0042】次に、推論条件選択部56において、複数の種別のキーワードから推論に使用する適切な商品と動詞を選択する組合せの例を挙げると、以下の通りである。

- (1) 商品キーワード中の商品と動詞
- (2) 商品キーワード中の商品と、動詞キーワード
- (3) 職業キーワード中の商品と動詞
- (4) 商品キーワード中の商品と、職業キーワード中の動詞

組合せが複数ある場合には、それぞれの組合せに対し商品と動詞による職業の推論を行うが、もし商品と動詞の関連がないならば商品と動詞の組合せからは職業は推論されない。以下に具体例を述べる。

【0043】＜入力文「布団を扱っているレンタルショップ」の場合＞商品キーワード「布団」、動詞キーワード「扱う」、職業キーワード「レンタルショップ」なので、(2)「布団」＋「扱う」と(4)「布団」＋「レンタルする」の組合せが生成される。このとき、商品と動詞による職業の推論によって、「布団」＋「扱う」からは対応する職業が推論されず、「布団」＋「レンタルする」から職業「貸し布団」が推論され、最終的には「貸し布団」が推論される。

【0044】＜入力文「貸しCDを扱っている店」の場合＞商品キーワード「貸しCD」、動詞キーワード「扱う」なので、(1)「CD」＋「貸す」、(2)「CD」＋「扱う」の組合せが生成される。このとき、商品と動詞による職業の推論によって、「CD」＋「扱う」からは対応する職業が推論されず、「CD」＋「貸す」から職業「CDレンタル店」が推論され、最終的には「CDレンタル店」が推論される。

【0045】＜入力文「ピアノが習える幼児教室」の場合＞商品キーワード「ピアノ」、動詞キーワード「習う」、職業キーワード「幼児教室」なので、(2)「ピアノ」＋「習う」、(3)「幼児」＋「教える」、(4)「ピアノ」＋「教える」の組合せが生成される。

【0046】次に、図8を用いて、プランニング知識テーブルの利用例について説明する。プランニング知識テーブルとは、動詞のとるべき商品の種別と動詞の組合せと、その組合せから変換可能な商品種別と動詞の組合せとを1レコードとするテーブルである。（例：（資格）取る→（事）習う）。複合語解析部54では、商品キーワード、動詞キーワード、職業キーワードから得られた商品と動詞に対し、その商品の商品種別と動詞から動詞をプランニング知識テーブル47を参照して別の動詞に変換し、さらに事象変換テーブル46を参照して、商品を新しく得られた動詞に対応した商品に変換する。これによって、商品と動詞の組合せに対する職業分類がない時、その商品と動詞の組合せを、その意味を損ねないような商品と動詞の組合せに変換して関連を再検索することができる。以下に具体例を述べる。

【0047】＜入力文「自動車免許を取りたい」の場合＞商品キーワード「自動車免許」、動詞キーワード「取る」が抽出されるが、このとき、「自動車免許」＋「取る」からは対応する職業が推論されない。そこで、プランニング知識テーブル47を参照して、動詞「取る」（資格）を「習う」（事）に変換し、さらに事象変換テーブル46を参照して、商品「自動車免許」（資格）を、「運転」（事）に変換する。この変換後の商品「運転」＋動詞「習う」によって、職業「自動車教習所」を推論することが可能になる。

【0048】

【発明の効果】以上説明したように、本発明では、入力された複合語のキーワードを前方削除したり、語尾やサ変名詞を削除したり、商品と動詞に分離したりして、複合語の主要な部分を取りだし、シソーラスを再検索するため、シソーラスには1～2語程度の要素からなる単語を登録しておくだけで、多様な複合語の入力キーワードに対応することが可能となる。

【0049】また、キーワードから抽出した商品と動詞に対し、商品を動詞に対応した種別に交換したり、動詞を別の動詞に変換するので、ユーザの意図に即した商品と動詞の組合せを得ることができる。

【0050】更に、複数の種類のキーワードが入力された場合でも、それらのキーワードから推論に必要な商品と動詞の組合せを選択して推論するので、正確に検索することができる。

【図面の簡単な説明】

【図1】本発明における情報検索装置の一実施例の全体構成図である。

50 【図2】知識ベース記憶装置に記憶される知識の一覧を

示す図である。

【図3】複合語解析部における、商品キーワードから商品と動詞を抽出する処理の一実施例を示す図である。

【図4】複合語解析部における、商品キーワードから商品と動詞を抽出するときの商品キーワードの遷移過程を示す図である。

【図5】複合語解析部における、職業キーワードから商品と動詞を抽出する処理の実施例を示す図である。

【図6】複合語解析部における、職業キーワードから動詞を抽出する処理の実施例を示す図である。

【図7】商品キーワードまたは職業キーワードから抽出した商品と動詞の組合せに対して、商品を動詞に対応し*

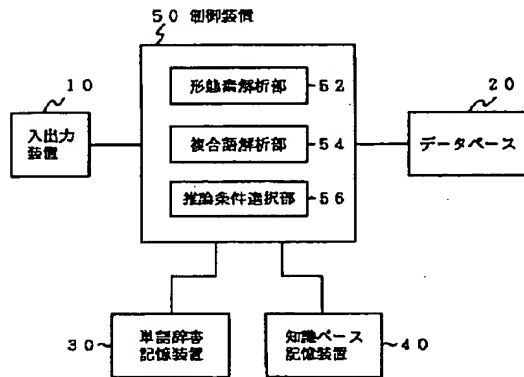
*た変換する商品最適変換の実施例を示す図である。

【図8】商品と動詞の組合せを、他の商品と動詞の組合せに変換する事象最適変換の一実施例を示す図である。

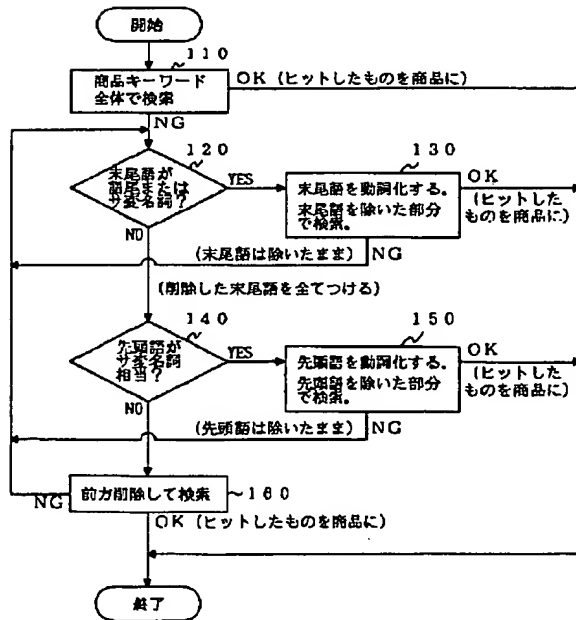
【符号の説明】

- 10 入出力装置
- 20 データベース
- 30 単語辞書記憶装置
- 40 知識ベース記憶装置
- 50 データベース
- 52 形態素解析部
- 54 複合語解析部
- 56 推論条件選択部

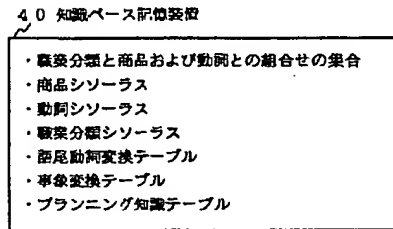
【図1】



【図3】

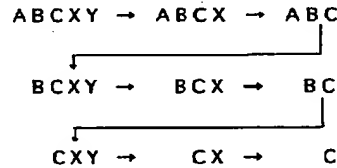


【図2】

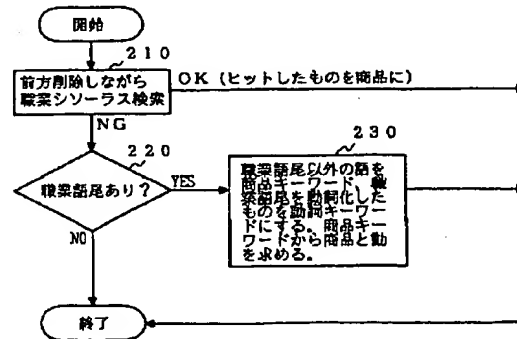


【図4】

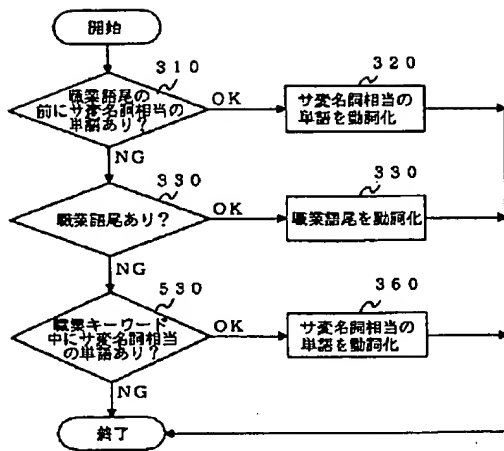
A B C : 通常の単語, X Y : 節尾, サ変



【図5】

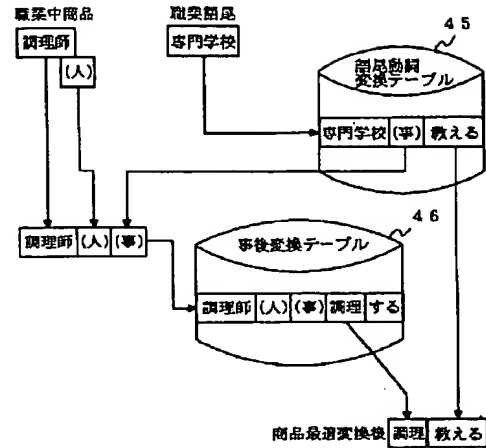


【図6】



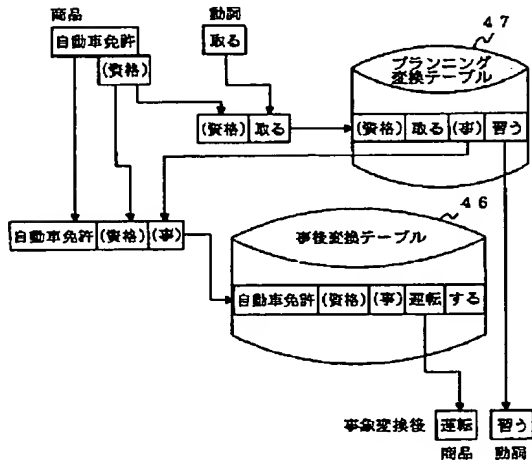
【図7】

職業キーワード「調理師/専門学校」



【図8】

入力文「自動車免許を取りたい」



【公報種別】特許法第17条の2の規定による補正の掲載
 【部門区分】第6部門第3区分
 【発行日】平成14年1月25日(2002.1.25)

【公開番号】特開平10-78969
 【公開日】平成10年3月24日(1998.3.24)
 【年通号数】公開特許公報10-790
 【出願番号】特願平8-233485
 【国際特許分類第7版】

G06F 17/30

【F I】

G06F 15/403 330 C
 320 D

【手続補正書】

【提出日】平成13年6月13日(2001.6.13)

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正内容】

【書類名】 明細書

【発明の名称】 情報検索装置

【特許請求の範囲】

【請求項1】 自然文で問合せ文を入力し、入力条件に合致する情報を出力する情報検索装置において、
 名義、商品、職業等を含む情報を格納するデータベースと、
 前記データベース中の名義、商品、職業等のキーワードを構成する単語と各単語の属性を対の形で記憶しておく単語辞書記憶装置と、
 前記データベースの職業分類と、該職業分類の業務対象を商品と動詞の組合せで記憶し、さらに商品や動詞シソーラスを記憶する知識ベース記憶装置と、
 複合語のキーワードが入力されたとき該複合語を解析し、商品と動詞に分離する複合語解析手段と、
 複数の種類のキーワードから推論に使用する商品と動詞を選択する推論条件選択手段を具備し、
 前記複合語解析手段は、複合語のキーワードを前方削除することにより、前記知識ベース記憶装置中のシソーラスを再検索することを特徴とする情報検索装置。

【請求項2】 前記複合語解析手段は、商品キーワードが語尾やサ変名詞を含む場合、該語尾やサ変名詞を削除して前記知識ベース記憶装置中の商品シソーラスを再検索し、削除した語尾やサ変名詞を動詞化して前記知識ベース記憶装置中の動詞シソーラスを検索することを特徴とする請求項1記載の情報検索装置。

【請求項3】 前記複合語解析手段は、職業キーワードが複合語の場合、該職業キーワードを商品部とサ変名詞

・職業語尾等の動詞部に分離し、商品部で商品シソーラスを検索し、動詞部は動詞化することにより動詞シソーラスを検索することを特徴とする請求項1記載の情報検索装置。

【請求項4】 前記知識ベース記憶装置は商品を別の商品と動詞の組合せで表わした事象変換テーブルを具備し、前記複合語解析手段は、該事象変換テーブルを用いて、商品キーワードまたは職業キーワードから抽出した商品と動詞の組合せに対して、商品を動詞に対応した商品に変換することを特徴とする請求項1記載の情報検索装置。

【請求項5】 前記知識ベース記憶装置は、商品を別の商品と動詞の組合せで表わした事象変換テーブル、及び、動詞のとるべき商品の種別と動詞の組合せと、当該組合せから変換可能な商品種別と動詞の組合せを表わすブランチング知識テーブルを具備し、
 前記複合語解析手段は、商品キーワード、動詞キーワード、職業キーワードから得られた商品と動詞に対し、その商品の商品種別と動詞から動詞を前記ブランチング知識テーブルを参照して別の動詞に変換し、さらに前記事象変換テーブルを参照して、商品を新しく得られた動詞に対応した商品に変換することを特徴とする請求項1記載の情報検索装置。

【請求項6】 前記推論条件選択手段は、商品キーワードと職業キーワード及び動詞キーワードが入力されたときは、入力された商品キーワードと商品キーワード及び職業キーワード中の商品部、商品キーワードと職業キーワード中の動詞部と動詞キーワードから、商品と動詞を組合せてそれぞれ推論することを特徴とする請求項1記載の情報検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報検索装置に係り、詳しくは、イエローページサービス等の職業情報案内サービスにおいて、問合せがデータベースに登録され

ている職業名と異なっていたり、問合せ中に職業名が含まれていなくても検索を可能とする情報検索装置に関するものである。

【0002】

【従来の技術】従来の自然語による情報検索では、検索に使用するためのキーワードが一語からなる単語または2語程度からなる複合語を対象としており、複合語のままシソーラスに登録し、検索処理を行っている。また、文献検索の分野では、複合語のキーワードを単語に分解し、それぞれの単語で検索を行っている。

【0003】一方、電話帳等の職業情報案内サービスにおいて、例えば特開昭64-29929号公報に記載のように、問合せ文中に職業名がなくとも問合せ文中の商品情報と動詞情報により職業を推論して、データベースを検索可能とするものもある。

【0004】

【発明が解決しようとする課題】あらゆる複合語のキーワードをシソーラスに登録することは、複合語を構成する単語の組合せの数が膨大になるので、非現実的である。そこで、シソーラスに登録される単語は、1～2語程度の単語に限定される。従来の自然語による情報検索では、入力されたキーワードをそのまま使用してシソーラスを検索するので、検索に使用できるキーワードに限定があり、ユーザの真の意図に合致する検索結果が得られないという問題があった。

【0005】また、複合語を単語に分解し、それぞれの単語で検索する方法では、問合せと無関係な結果が得られるという問題がある。特に職業情報案内の分野では適切な職業のみ推論することが重要である。例えば複合語中の単語の役割を意識しないと、「競馬情報」という入力で「情報」から「情報科学」を検索してしまうという問題がある。さらに、「布団を扱うレンタルショップ」の様に動詞キーワードが職業（「レンタルショップ」）の中に含まれている場合がある。この場合は動詞キーワード「扱う」から推論に必要な動詞「レンタルする」を導き出すことが出来ない。

【0006】更に、特開昭64-29929号公報に記載の方法は、単純な意味解析によって問合せ文中から商品情報と動詞情報を得るものであるため、多様な複合語の入力キーワードに対応できない問題があった。

【0007】本発明の目的は、シソーラスには1～2語程度の要素からなる単語を登録しておくだけで、多様な複合語の入力キーワードに対応することが可能な情報検索装置を提供することにある。

【0008】本発明の他の目的は、複数の種類のキーワードが入力された場合でも、それらのキーワードから推論に必要な商品と動詞の組合せを選択して推論することで、正確に検索することが可能な情報検索装置を提供することにある。

【0009】

【課題を解決するための手段】複合語の商品キーワードが入力された時、語尾（複合語の末尾にあって、複合語の中では主要な意味をもたない単語）やサ変名詞（「する」をつけて動詞になる名詞。例：「予想」）が含まれるときは、語尾やサ変名詞を動詞化し、商品キーワードからは語尾やサ変名詞は削除する。商品キーワードに語尾やサ変名詞が含まれないときは、商品キーワードを前方削除してキーワードを再構成する。

【0010】複合語の職業キーワードが入力されたときは、職業キーワードを商品部分と、語尾サ変等の動詞部分に分離し、商品部分から商品キーワード、動詞部分から動詞キーワードを再構成する。

【0011】このようにして、商品キーワードと動詞キーワードを作り直し、知識ベース記憶装置中のシソーラスを検索する。商品キーワードまたは動詞キーワードからシソーラス上での商品と動詞が決定すれば、商品と動詞に対応した商品に変換する。

【0012】以上の過程から決定した商品と動詞を用いて、データベースの職業分類と、該職業分類の業務対象を商品と動詞の組合せで記憶する知識ベースを検索し、商品と動詞から職業を推論する。また、商品キーワードと職業キーワード及び動詞キーワードが入力されたときは、複合語解析装置が出力した、商品キーワード・職業キーワード中の商品、動詞と動詞キーワードを組合せて推論する。さらに、商品と動詞の組合せに対する職業分類がない時、その商品と動詞の組合せを、その意味を損ねないような商品と動詞の組合せに変換して知識ベースを再検索する。

【0013】一般に日本語では複合語を構成する単語の内、後方の部分に重要な意味をもっているものが多いため、前方削除によりシソーラス中の単語を検索することにより適切な商品を検索できる。一方、日本語では、複合語の末尾にくる単語があまり意味のない単語であったり、サ変名詞のように動詞化できることもあり、そのような末尾語やサ変名詞を動詞化し、末尾語・サ変名詞を除いた複合語中の意味的に重要な部分でシソーラス中の単語を検索することにより適切な商品を検索できる。

【0014】商品キーワードや職業キーワードから抽出した商品と動詞は、商品が動詞のとる格に設定できない場合があるので、商品と動詞の格に対応した商品に変換することにより適切な商品と動詞の組を得ることができる。また、複数の種類のキーワードがあった場合でも、それらのキーワードから推論に必要な商品と動詞の組合せを選択して推論するので、検索結果に誤りをなくすることができる。

【0015】

【発明の実施の形態】以下、本発明の一実施例について図面により説明する。

【0016】図1は本発明における情報検索装置の一実施例の全体構成図を示す。図において、10は問合せ文

を入力し、検索結果を出力する入出力装置、20は名義、商品、職業等を含む情報を格納するデータベース、30は名義、商品、職業等のキーワードを構成する単語と各単語の属性を対の形で記憶しておく単語辞書記憶装置、40は種々の知識を記憶する知識ベース記憶装置、50は全体を制御する制御装置である。ここで、制御装置50は、本発明に関連する構成として、入力された問合せ文を単語辞書記憶装置30を用いて単語単位に分解する形態素解析部52、複合語の商品キーワードや職業キーワードなどを解析し、商品と動詞に分離する複合語解析部54、複数の種類のキーワードから推論に使用する商品と動詞を選択する推論条件選択部56を有する。

【0017】図2は、知識ベース記憶装置40に記憶される知識の一覧である。ここで、データベースの職業と、該職業分類の業務対象を商品と動詞の組合せで表わした集合（職業分類＝商品＋動詞）は、最終的に商品と動詞から職業を推論するときに使用される。語尾動詞交換テーブルは職業語尾と動詞の対応を表わしたテーブルである（例：音楽教室の教室→教える）。事象交換テーブルは商品を別の商品と動詞の組合せで表わしたテーブルである（例：調理師→調理＋する）。プランニング知識テーブルは、動詞のとるべき商品の種別と動詞の組合せと、その組合せから交換可能な商品種別と動詞の組合せを表わしたテーブルである（例：（資格）取る→（事）習う）。

【0018】以下では、入出力装置10から入力された問合せ文は、すでに形態素解析部52において単語辞書記憶装置30を用いて単語単位に分解されているとして、特に複合語解析部54の動作について詳述する。

【0019】まず、図3を用いて、複合語解析部54が複合語の商品キーワードからシンソーラス中の商品と動詞を求める処理例を説明する。

【0020】ステップ110；

まず、商品キーワード全体で、知識ベース記憶装置40の商品シンソーラスを検索する。ここで、検索に成功すれば、それを商品として終了する。失敗した場合、ステップ120に進む。

【0021】ステップ120；

末尾語が語尾またはサ変名詞か判断する。語尾またはサ変名詞ならば、ステップ130に進む。それ以外ならば、ステップ140に進む。例えば、「競馬予想」「競馬情報」「ピアノの先生」では、それぞれ「予想」はサ変名詞、「情報」「先生」は語尾なので、ステップ130に進む。また、「家庭菜園」の末尾語「菜園」は語尾でもサ変名詞でもないので、ステップ140に進む。

【0022】ステップ130；

末尾語が語尾の場合は、知識ベース記憶装置40の語尾動詞交換テーブルを参照し、動詞に変換する。サ変名詞の場合は、末尾に「する」を付加して動詞にする。例えば、「予想」は「予想する」、「先生」は「教える」に

する。「情報」は動詞には対応しないので、動詞なしとする。その後、末尾語を除いた部分で知識ベース記憶装置40の商品シンソーラスを検索する。例えば、「競馬予想」の場合は「競馬」で商品シンソーラスを検索する。そして、成功すれば、それを商品とし終了する。失敗した場合は、末尾語を除いたままステップ120に戻り、次に、末尾語を除いた部分の末尾語を判定する。

【0023】ステップ140；

それまで削除した末尾語を全て戻した後、先頭語がサ変名詞相当かどうか判断する。そして、サ変名詞相当ならば、ステップ150に進む。そうでなければ、ステップ160に進む。例えば、「レンタルCD」「貸し衣装」の「レンタル」や「貸し」はサ変名詞相当で、それぞれ「レンタルする」や「貸す」に動詞化できるので、ステップ150に進む。

【0024】ステップ150；

先頭語を動詞化する。その後、先頭語を除いた部分で知識ベース記憶装置40の商品シンソーラスを検索する。例えば、「レンタルCD」の場合は「CD」で検索する。そして、成功すれば、それを商品とし終了する。失敗した場合は、先頭語を除いたままステップ120に戻る。

【0025】ステップ160；

先頭語を削除して、知識ベース記憶装置40の商品シンソーラスを検索する。例えば、「ビジネス英語」の場合「ビジネス」を削除し、「英語」で検索する。そして、検索に成功すれば、それを商品として、終了する。失敗した場合は、先頭語を削除したままステップ120に戻る。

【0026】図4に、複合語の商品キーワードから商品と動詞を求める処理における商品キーワードの遷移過程を示す。即ち、まず、末尾語にくる語尾やサ変名詞を削除していき、末尾語の語尾やサ変名詞がなくなったならば、一旦削除した語尾・サ変名詞を全てつけ戻した上で、次に、先頭語を削除し、再び末尾の語尾・サ変名詞を削除していく。図4は、最終的に、ステップ130で単語「C」が検索に成功したことを示している。

【0027】次に、図5を用いて、複合語解析部54が複合語の職業キーワードからシンソーラス中の商品と、動詞を求める処理例を説明する。

【0028】ステップ210；

複合語の職業キーワードを一語ずつ前方削除しながら、知識ベース記憶装置40の職業シンソーラスを検索する。例えば、「温泉旅館」の場合、「温泉」を削除して、「旅館」で検索する。そして、検索に成功すれば、それを商品として、終了する。失敗した場合は、ステップ220に進む。

【0029】ステップ220；

職業語尾（職業相当の語尾）があるかどうか判断する。そして、職業語尾があるならば、ステップ230に進む。なければ終了する。例えば、「音楽教室」では「教

室」は職業語尾なので、ステップ230に進む。「レストラン」は職業語尾でないので終了する。

【0030】ステップ230；

まず、職業語尾以外の語を商品キーワードとし、職業語尾は知識ベース記憶装置40の語尾変換テーブルを参照して動詞に変換し動詞キーワードとする。例えば、「音楽教室」では「音楽」を商品キーワードに、「教室」を「教える」に動詞化して動詞キーワードにする。次に、商品キーワードについて、図3で述べた処理を行い、シソーラス中の商品と動詞を求める。

【0031】次に、図6を用いて、複合語解析部54が職業キーワードから動詞を抽出する処理例を説明する。

【0032】ステップ310；

職業語尾の直前の単語がサ変名詞相当の単語かどうか判断する。職業語尾がない場合は、末尾の語についてサ変名詞相当かどうか判断する。そして、サ変名詞相当の単語ならば、ステップ320に進む。そうでなければ、ステップ330に進む。例えば、「ビデオレンタル店」では、「レンタル」はサ変名詞なので、ステップ320に進む。

【0033】ステップ320；

サ変名詞相当の単語を動詞化して終了する。

【0034】ステップ330；

職業語尾があるかどうか判断する。あれば、ステップ340に進む。なければ、ステップ350に進む。例えば、「音楽教室」では「教室」は職業語尾なので、ステップ340に進む。

【0035】ステップ340；

職業語尾を知識ベース記憶装置40の語尾動詞変換テーブルを参照して動詞に変換して終了する。

【0036】ステップ350；

職業キーワードの先頭の単語からサ変名詞相当の単語かどうか調べていく。サ変名詞相当の単語が見つければ、ステップ360に進む。見つからなければ、終了する。例えば、「貸しビデオ店」では、「貸し」はサ変名詞相当の単語なので、ステップ360に進む。

【0037】ステップ360；サ変名詞相当の単語を動詞化して終了する。例えば、サ変名詞相当の単語「貸し」は「貸す」に動詞化して終了する。

【0038】次に、図7を用いて、複合語解析部54が商品キーワードまたは職業キーワードから抽出した商品と動詞の組合せに対して、商品を動詞に対応した商品に変換する商品最適変換の処理例を説明する。

【0039】知識ベース記憶装置40においては、各商品には物、事、人等の商品種別が割り当てられている。例えば、「布団」、「競馬」、「調理師」の商品種別はそれぞれ「物」、「事」、「人」である。一方、動詞にはその動詞がとるべき商品の種別が定まっている。例えば、「教える」がとるべき商品の種別は「事」である。

【0040】キーワードから商品と動詞を抽出した場

合、動詞と商品種別との不整合がある場合がある。このような場合、複合語解析部54では、図7のように、知識ベース記憶装置40の商品を動詞に対応した商品種別の商品に対応づけしている事象変換テーブル46を用いて商品の最適変換を行う。

【0041】例えば、職業キーワードが「調理師／専門学校」の場合、職業中商品として「調理師（商品種別：人）」が抽出され、職業中動詞として「教える」（専門学校→教える）が語尾動詞変換テーブル45により抽出される。しかしながら、「教える」がとるべき商品種別は「事」であるので、「調理師（人）」が「する事」は何かを、事象変換テーブル46を参照して変換すると「調理（事）」という知識が得られる。そこで、商品を「調理」に変換する。これによって、職業分類および商品と動詞との組合せ集合により、商品「調理」と動詞「教える」から職業「料理学校」を推論することができる。

【0042】次に、推論条件選択部56において、複数の種別のキーワードから推論に使用する適切な商品と動詞を選択する組合せの例を挙げると、以下の通りである。

- (1) 商品キーワード中の商品と動詞
- (2) 商品キーワード中の商品と、動詞キーワード
- (3) 職業キーワード中の商品と動詞
- (4) 商品キーワード中の商品と、職業キーワード中の動詞

組合せが複数ある場合には、それぞれの組合せに対し商品と動詞による職業の推論を行うが、もし商品と動詞の関連がないならば商品と動詞の組合せからは職業は推論されない。以下に具体例を述べる。

【0043】＜入力文「布団を扱っているレンタルショップ」の場合＞

商品キーワード「布団」、動詞キーワード「扱う」、職業キーワード「レンタルショップ」なので、(2)「布団」＋「扱う」と(4)「布団」＋「レンタルする」の組合せが生成される。このとき、商品と動詞による職業の推論によって、「布団」＋「扱う」からは対応する職業が推論されず、「布団」＋「レンタルする」から職業「貸し布団」が推論され、最終的には「貸し布団」が推論される。

【0044】＜入力文「貸しCDを扱っている店」の場合＞

商品キーワード「貸しCD」、動詞キーワード「扱う」なので、(1)「CD」＋「貸す」、(2)「CD」＋「扱う」の組合せが生成される。このとき、商品と動詞による職業の推論によって、「CD」＋「扱う」からは対応する職業が推論されず、「CD」＋「貸す」から職業「CDレンタル店」が推論され、最終的には「CDレンタル店」が推論される。

【0045】＜入力文「ピアノが習える幼児教室」の場

合>

商品キーワード「ピアノ」、動詞キーワード「習う」、職業キーワード「幼児教室」なので、(2)「ピアノ」+「習う」、(3)「幼児」+「教える」、(4)「ピアノ」+「教える」の組合せが生成される。

【0046】次に、図8を用いて、ブランニング知識テーブルの利用例について説明する。ブランニング知識テーブルとは、動詞のとるべき商品の種別と動詞の組合せと、その組合せから変換可能な商品種別と動詞の組合せとを1レコードとするテーブルである。(例：(資格)取る→(事)習う)。複合語解析部54では、商品キーワード、動詞キーワード、職業キーワードから得られた商品と動詞に対し、その商品の商品種別と動詞から動詞をブランニング知識テーブル47を参照して別の動詞に変換し、さらに事象変換テーブル46を参照して、商品を新しく得られた動詞に対応した商品に変換する。これによって、商品と動詞の組合せに対する職業分類がない時、その商品と動詞の組合せを、その意味を損ねないような商品と動詞の組合せに変換して商品と動詞の組合せ集合を再検索することができる。以下に具体例を述べる。

【0047】<入力文「自動車免許を取りたい」の場合>

商品キーワード「自動車免許」、動詞キーワード「取る」が抽出されるが、このとき、「自動車免許」+「取る」からは対応する職業が推論されない。そこで、ブランニング知識テーブル47を参照して、動詞「取る」(資格)を「習う」(事)に変換し、さらに事象変換テーブル46を参照して、商品「自動車免許」(資格)を、「運転」(事)に変換する。この変換後の商品「運転」+動詞「習う」によって、職業「自動車教習所」を推論することが可能になる。

【0048】

【発明の効果】以上説明したように、本発明では、入力された複合語のキーワードを前方削除したり、語尾やサ変名詞を削除したり、商品と動詞に分離したりして、複合語の主要な部分を取りだし、シソーラスを再検索するため、シソーラスには1〜2語程度の要素からなる単語を登録しておくだけで、多様な複合語の入力キーワードに対応することが可能となる。

【0049】また、キーワードから抽出した商品と動詞

に対し、商品を動詞に対応した種別に変換したり、動詞を別の動詞に変換するので、ユーザの意図に即した商品と動詞の組合せを得ることができる。

【0050】更に、複数の種類のキーワードが入力された場合でも、それらのキーワードから推論に必要な商品と動詞の組合せを選択して推論するので、正確に検索することができる。

【図面の簡単な説明】

【図1】本発明における情報検索装置の一実施例の全体構成図である。

【図2】知識ベース記憶装置に記憶される知識の一覧を示す図である。

【図3】複合語解析部における、商品キーワードから商品と動詞を抽出する処理の一実施例を示す図である。

【図4】複合語解析部における、商品キーワードから商品と動詞を抽出するときの商品キーワードの遷移過程を示す図である。

【図5】複合語解析部における、職業キーワードから商品と動詞を抽出する処理の実施例を示す図である。

【図6】複合語解析部における、職業キーワードから動詞を抽出する処理の実施例を示す図である。

【図7】商品キーワードまたは職業キーワードから抽出した商品と動詞の組合せに対して、商品を動詞に対応した変換する商品最適変換の実施例を示す図である。

【図8】商品と動詞の組合せを、他の商品と動詞の組合せに変換する事象最適変換の一実施例を示す図である。

【符号の説明】

- 10 入出力装置
- 20 データベース
- 30 単語辞書記憶装置
- 40 知識ベース記憶装置
- 50 制御装置
- 52 形態素解析部
- 54 複合語解析部
- 56 推論条件選択部

【手続補正2】

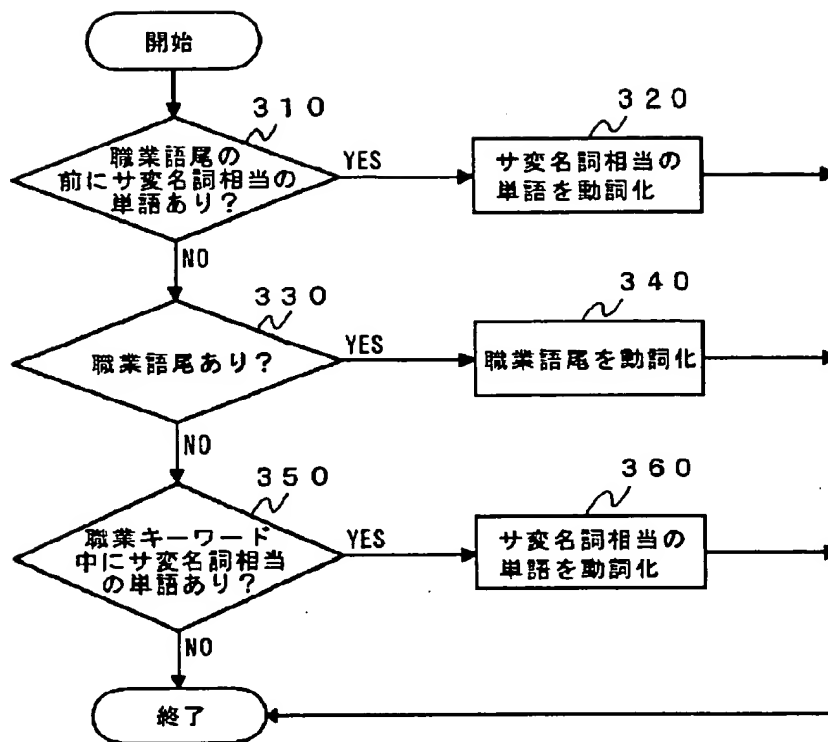
【補正対象書類名】図面

【補正対象項目名】図6

【補正方法】変更

【補正内容】

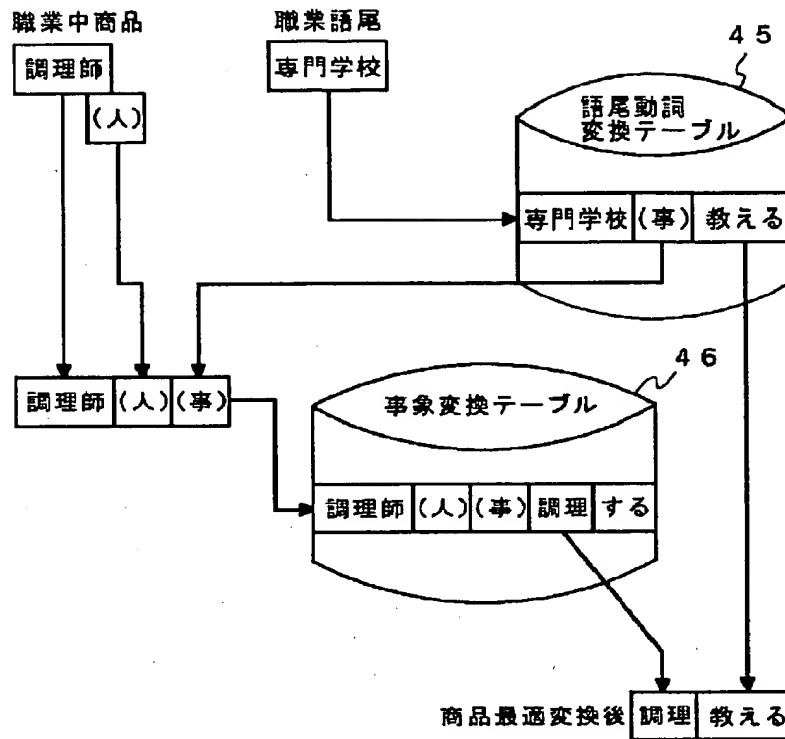
【図6】



【手続補正3】
 【補正対象書類名】図面
 【補正対象項目名】図7

【補正方法】変更
 【補正内容】
 【図7】

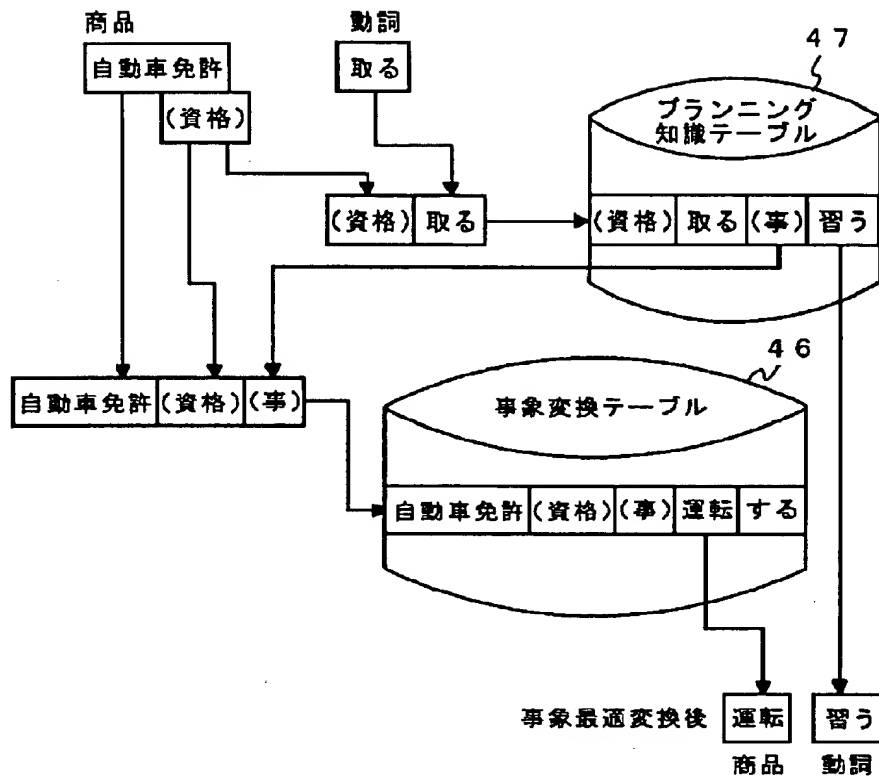
職業キーワード「調理師／専門学校」



【手続補正4】
 【補正対象書類名】図面
 【補正対象項目名】図8

【補正方法】変更
 【補正内容】
 【図8】

入力文「自動車免許を取りたい」



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.